International Journal of Internet, Broadcasting and Communication Vol.13 No.2 7-13 (2021)
http://dx.doi.org/10.7236/IJIBC.2021.13.2.7

IJIBC 21-2-2

Image Enhanced Machine Vision System for Smart Factory

ByungJoo Kim

Professor, Department of Electrical and Electronic Engineering, Youngsan University, Korea
bjkim@ysu.ac.kr

Abstract

Machine vision is a technology that helps the computer as if a person recognizes and determines things. In recent years, as advanced technologies such as optical systems, artificial intelligence and big data advanced in conventional machine vision system became more accurate quality inspection and it increases the manufacturing efficiency. In machine vision systems using deep learning, the image quality of the input image is very important. However, most images obtained in the industrial field for quality inspection typically contain noise. This noise is a major factor in the performance of the machine vision system. Therefore, in order to improve the performance of the machine vision system, it is necessary to eliminate the noise of the image. There are lots of research being done to remove noise from the image. In this paper, we propose an autoencoder based machine vision system to eliminate noise in the image. Through experiment proposed model showed better performance compared to the basic autoencoder model in denoising and image reconstruction capability for MNIST and fashion MNIST data sets.

Keywords: *Image Enhancing, Convolutional Neural Network, Autoencoder, Machine Vision.*

1. Introduction

In recent years, the life cycle of products has shortened and consumer requirements have diversified, increasing the requirements for customized individual production. As the economic structure shifted from manufacturing to the center of the service industry, including information and communication technology (ICT), traditional manufacturing demanded innovation, which led to the emergence of a new smart factory. Smart factories are advanced intelligent factories that integrate the entire process from product planning to design, production, distribution and sales into information and communication technology (ICT) to produce customized products at minimal cost and time. Smart factories are flexible and intelligent factories that connect, collect, and analyze data with technologies such as artificial intelligence, big data, the Internet of Things (IoT), and wireless communication, and are separated from factory automation, which seeks to unmanned and automate production processes using equipment such as computers and robots. Because smart factories differ in their characteristics and operating methods by manufacturing, companies should establish and introduce smart factory construction plans step-by-step in a direction that reflects the unique value of the enterprise. Among the many areas required to build a smart factory, automation of the manufacturing process remains to

Manuscript Received: January. 13, 2021 / Revised: January. 22, 2021 / Accepted: January. 27, 2021
Corresponding Author: bjkim@ysu.ac.kr
Tel: +82-055-380-9447
Professor, Department of Electrical & Electronic Engineering, Youngsan University, Korea

be addressed. Advances in technology that can automate these areas will help us reach the ultimate smart factory. Machine vision is a technique that quickly and precisely replaces the task of judging the human eye through a system consisting of software such as processing and analyzing images in the manufacturing process. In other words, machine vision is a technology that helps the computer as if a person recognizes and determines things. In recent years, as advanced technologies such as optical systems, artificial intelligence and big data advanced in conventional machine vision system became more accurate quality inspection and it increases the manufacturing efficiency. With the appearance of deep learning, the practicality of machine vision has been enhanced, and the realm of artificial intelligence has expanded. In machine vision systems using deep learning, the image quality of the input image is very important. However, most images obtained in the industrial field for quality inspection typically contain noise. This noise is a major factor in the performance of the machine vision system. Therefore, in order to improve the performance of the machine vision system, it is necessary to eliminate the noise of the image. There is a lot of research being done to remove noise from the image [1][2][3][4]. In this paper, an autoencoder is used to eliminate noise in the image [5]. We want to use this as input to the machine vision system to improve the performance of the machine vision system. Paper is organized as follows. Section 2 briefly outlines the autoencoder algorithm and denoising in image using basic autoencoder. In section 3 proposed model is presented and experimental results of the proposed system will be described in section 4. Finally, the analysis of the experimental results and the future research will be described in section 5.

2. Autoencoder

An autoencoder is a neural network that simply copies input into output. This may seem like a very simple neural network, but it can create a variety of neural networks by constraining the network architecture. For example, there may be a variety of autoencoders, such as compressing the data by reducing the number of neurons in the hidden layer or learning the network to reconstruct the original input data. The architecture of the autoencode used to reconstruction the input image is shown in Figure 1. As shown in Figure 1, the autoecoder largely consists of encoders and decoders. The encoder learns how to encode the input into a compressed representation, and the decoder reconstruct the original input. One of the important concepts of autoencoder is the latent representation of input data. Compressing an input into a smaller dimension allows autoencoder to learn the latent representation of the input. These latent representation reflects only the most important features of the input data.

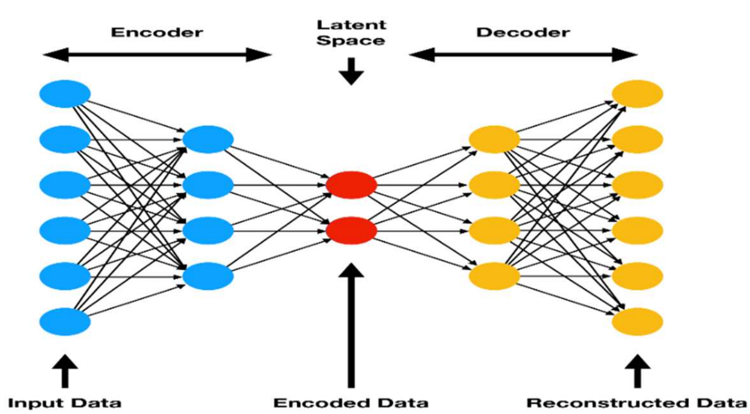


Figure 1. Architecture of autoencoder [6].

Autoencoder can be used as follows. First input data can be reduced in terms of dimensions. The second is that noise of input data can be removed. Because noise is not an important feature of the data, so it is necessary to remove the noise by latent representations.

2.1 Image Denoising with Basic Autoencoder

As mentioned earlier, autoencoder can be used to image denoising. Image noise is a phenomenon in which the brightness of pixels in an image changes randomly, usually due to sensors in digital cameras. Today, digital camera can shoot high-quality images, but image noise still occurs, especially in low-light environments. Recently, autoencoder which is one of the deep learning techniques was proposed to remove the noise in the images. The architecture of basic autoencode for image denoising is very simple. This method uses noisy images as input data and provides clean images with target answer to output layer, which makes autoencoder learns how to denoise. Figure 2. Shows the denoising process of autoencoder.

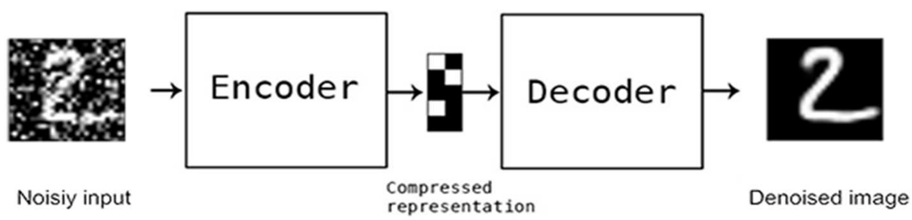


Figure 2. A denoising autoencoder process [7].

At first we build a basic autoencoder model. The model consists of three parts input, hidden and output layer. We use input as MNIST data set of handwritten digts [8]. Since each image consists of 28 pixels in width and length, the size of the input layer is 784. When building an autocorder the crucial part is determining the architecture of hidden layer. In this case we set 1 hidden layer and size of the hidden layer is 16. Activation function from input to hidden layer is relu. Output layer is fully connect and size of the output layer is 784. Activation function from hidden to output layer is sigmoid. Applying relu function to the hidden layer and the sigmoid function to the output layer generally results in much higher accuracy than before. The result of the experiment is shown in Figure 3. The results of the experiment show that the noise was well removed, but the input image could not be reconstructed properly. The results of this experiment show that the image reconstruction results are not satisfactory when using a single hidden layer. When using this basic autoencoder model in machine vision system for defective inspection, it is difficult to expect high-performance defective inspections.

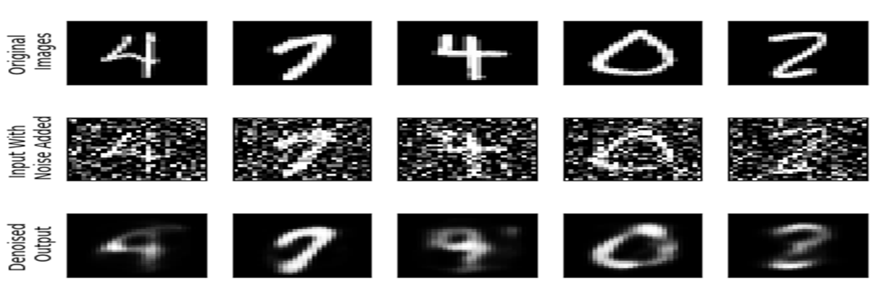


Figure 3. Image denoising by basic autoencoder.

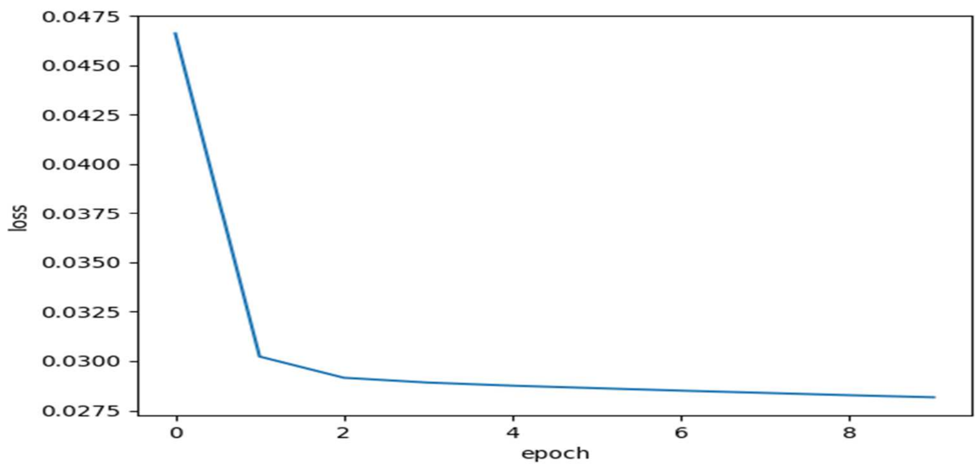


Figure 4. Reconstruction error value by epoch in basic autoencoder.

3. Proposed System

To improve the performance of reconstruction ability we will add convolutional layer to basic autoencoder model [9]. Many research show that feature extraction and image classification performance is improved when using deep convolution neural network method [10][11][12]. We adopt this idea and add the convolutional layer between input and hidden layer and also between hidden and output layer. There is no explicit rule setting the proper number of convolution layers. In this research we set the number of hidden layers as 2. Architecture of proposed model is shown in Figure 5.

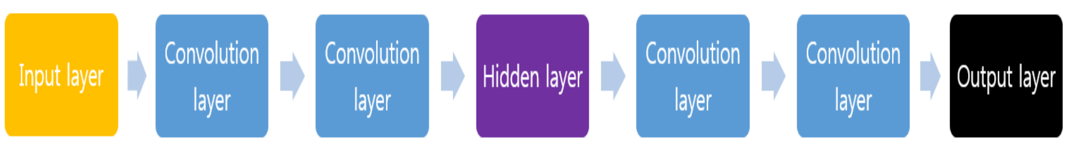


Figure 5. Architecture of proposed model.

4. Experiment

We will evaluate our proposed model on two data sets. The first data set is MNIST handwritten data set which we used in paragraph 2.1 and the other one is fashion MNIST data set [13].

4.1 MNIST Data Set

The number of filters in encoder part is as follows. The number of filters in first convolution layer is 16 and 8 on the second convolution layer respectively. On the other hand, the number of filters in decoder part is 8 and 16 in order. The size of the filter is generally size 3 in vertically and horizontally, so we apply it. We use Adam as optimization method [14]. The results performed through a 10 epoch are shown in Figure 6. As shown in Figure 6, the resolution of the reconstructed image as well as the noise removal performance is very high.

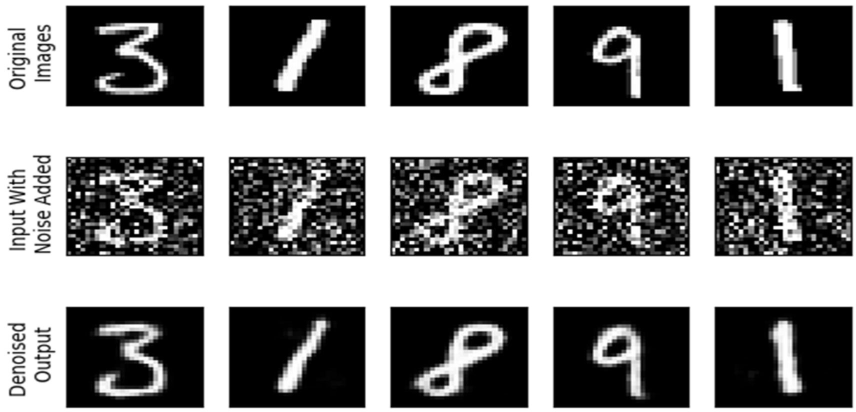
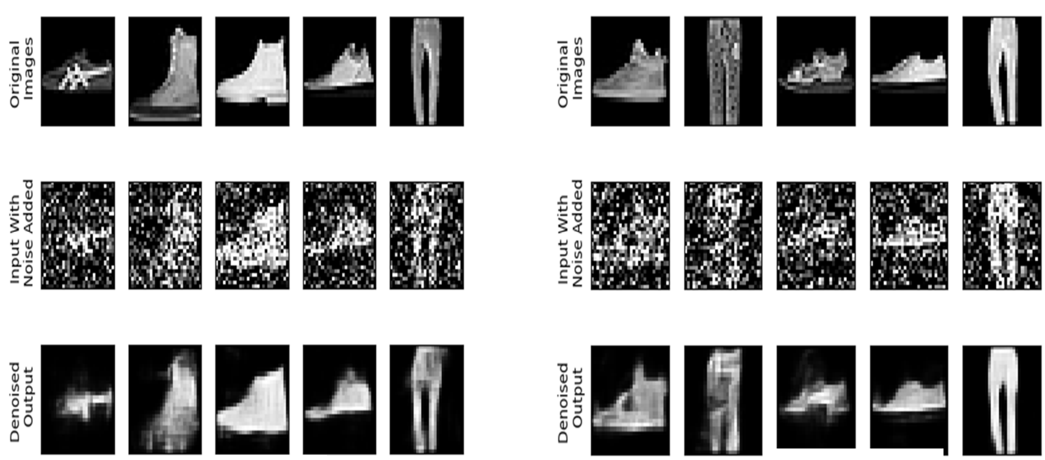


Figure 6. Experimental results by proposed model.

4.2 Fashion MNIST Data Set

We apply our proposed model to more complicated images. Images of fashion MNIST have more and more diverse forms, so image reconstruction is relatively more difficult than MNIST images. We added one more convolution layer because it applies for more complex data. The number of filters in encoding part is 32, 16 and 8 respectively in order. On the other hand, the number of filters in decoder part is 32, 16 and 8 in order. The size of the filter is same as on MNIST data set. We use RMSProp as optimizer. The results performed through a 10 epoch are shown in Figure 7. The resolution of the reconstructed image as well as the noise removal performance with 3 convolution layers which is shown in Figure 7(b) is better than using 2 convolution layers as shown in Figure 7(a).



(a) (b)

Figure 7. Image denoisng on fashion MNIST data set. Figure 7(a) is 2 convolution layers and Figure 7(b) is 3 convolution layers .

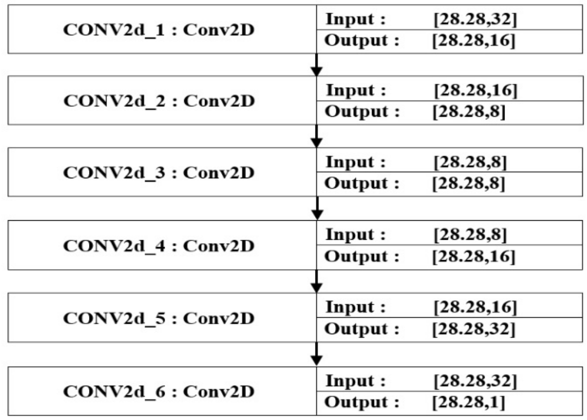


Figure 8. Model architecture and parameters in proposed model with 3 convolution layers.

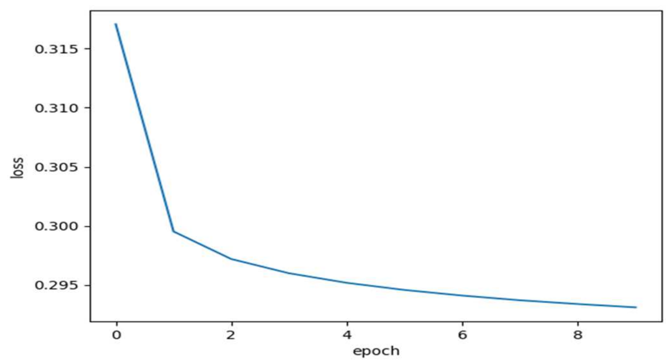


Figure 9. Reconstruction error value by epoch.

Figure 8 shows the model architecture and parameters in proposed model with 3 convolution layers and figure 9 shows the reconstruction error by proposed model with 3 convolution layers.

5. Experimental Results

Through experiment proposed model showed better performance compared to the basic autoencoder model in denoising and image reconstruction capability for MNIST and fashion MNIST data sets. The experimental results for each data set are as follows. In MNIST data set we propose a model which has 2 convolutional layers and Adam optimizer is used. Proposed model shows the excellent performance in image denoisng and reconstruction. We extend our experiment to more challenging data set. In fashion MNIST data set proposed model has 3 convolutional layers and RMSProp optimizer is used. Because images in fashion MNIST data set are more complex we set more convolutional layer so as to extract more useful feature in the image. Proposed convolutional autoencoder model also shows the excellent performance in image denoisng and reconstruction.

Models with three convolution layers have the disadvantage that it requires more training time than models with two convolution layers. Setting the appropriate number of convolution layers is not an easy task, which is a problem that must be solved through multiple attempts at the given task to be solved.

6. Conclusion

In this paper, we propose an autoencoder based machine vision system to eliminate noise in the image. To improve the performance of reconstruction ability we add convolutional layer to basic autoencoder model. Proposed model adds the convolutional layer between input and hidden layer and also between hidden and output layer. There is no explicit rule setting the proper number of convolution layers. In this research we set the number of hidden layers as 2. Through experiment proposed model showed better performance compared to the basic autoencoder model in denoising and image reconstruction capability for MNIST and more complicated fashion MNIST data sets. We will install the machine vision system to company Jade Solution. Future work is using the generated images from proposed model as input to machine vision system and testing the classification performance in manufacturing field. We will check the classification accuracy and will improve the system performance based on the field test.

Acknowledgement

This work was supported by Youngsan University Research Fund of (2020).

References

[1] T. Brooks, B. Mildenhall, T. Xue, J. Chen, D. Sharlet, and J. T. Barron, “Unprocessing images for learned raw denoising,” In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11036-11045, 2019. DOI: 10.1109/CVPR.2019.01129

[2] U. Dmitry, V. Andrea, and L. Victor, “Deep image prior,” In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 9446-9454, 2018. **DOI:** 10.1109/CVPR.2018.00984

[3] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, “Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising,” *IEEE transactions on image processing*, Vol. 26, No. 7, pp. 3142-3155, 2017. DOI: 10.1109/TIP.2017.2662206

[4] J. Kim, “Edge-Preserving and Adaptive Transmission Estimation for Effective Single Image Haze Removal,” *International Journal of Internet, Broadcasting and Communication*, Vol. 12, No 2, pp. 21-29, 2020. DOI : https://doi.org/10.7236/IJIBC.2020.12.2.21

[5] C. Doersch, “Tutorial on variational autoencoders,” arXiv preprint arXiv:1606.05908, 2016.

[6] *https://www.compthree.com/blog/autoencoder/*

[7] *https://www.pyimagesearch.com/2020/02/24/denoising-autoencoders-with-keras-tensorflow-and-deep-learning/*

[8] *http://yann.lecun.com/exdb/mnist/*

[9] T. N. Sainath, A. R. Mohamed, B. Kingsbury and B. Ramabhadran, “Deep convolutional neural networks for LVCSR,” In *2013 IEEE international conference on acoustics, speech and signal processing* , pp. 8614-8618, May 2013. DOI: 10.1109/ICASSP.2013.6639347

[10] A. Krizhevsky, I. Sutskever, and G. Hinton, “Imagenet classification with deep convolutional neural networks,” *Advances in neural information processing systems*, Vol. *25*, pp. 1097-1105, 2012. https://doi.org/10.1145/3065386

[11] T. N. Sainath, B. Kingsbury, A. R. Mohamed, G. E. Dahl, G. Saon, H. Soltau and B. Ramabhadran, “Improvements to deep convolutional neural networks for LVCSR,” In *2013 IEEE workshop on automatic speech recognition and understanding*, pp. 315-320, Dec. 2013. DOI: 10.1109/ASRU.2013.6707749

[12] A. Khan, A. Sohail, U. Zahoora and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” *Artificial Intelligence Review*, Vol. 53, No. 8, pp. 5455-5516, 2020. Doi : 10.1007/s10462-020-09825-6

[13] H. Xiao, K. Rasul, and R. Vollgraf, “Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms,” arXiv preprint arXiv:1708.07747, 2017.

[14] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” arXiv preprint arXiv:1412.6980, 2014.